

Conference Abstract

Reducing Dependence on Digital Biodiversity Data Silos Through Global Alignment and Collaboration

Gil Nelson[‡], Talia Karim[§], Rosemary Gillespie[|], Jose Fortes[¶], Douglas S Jones[‡]

[‡] Florida Museum of Natural History, University of Florida, Gainesville, FL, United States of America

[§] University of Colorado, Boulder, Boulder, CO, United States of America

[|] University of California Berkeley, Berkeley, United States of America

[¶] University of Florida, Gainesville, FL, United States of America

Corresponding author: Gil Nelson (gnelson@floridamuseum.ufl.edu)

Received: 03 Jul 2019 | Published: 10 Jul 2019

Citation: Nelson G, Karim T, Gillespie R, Fortes J, Jones DS (2019) Reducing Dependence on Digital Biodiversity Data Silos Through Global Alignment and Collaboration. Biodiversity Information Science and Standards 3: e37897. <https://doi.org/10.3897/biss.3.37897>

Abstract

Over the last decade, the Integrated Digitized Biocollections ([iDigBio](#)) organization and the Advancing the Digitization of Biodiversity Collections (ADBC) grant program, both funded by the US National Science Foundation (NSF), have made large strides in the aggregation of pre-existing siloed digital collections data as well as the new digitization of previously dark collections data across the United States. The impact of iDigBio leadership in community engagement (e.g., through discipline-specific workshops and webinars) and data mobilization (e.g., aggregation assistance, portal development) is widespread and with impact across all collection types and sizes. Moreover, the funding model for the ADBC program, which required the development of digitization-based Thematic Collection Networks (TCNs), facilitated engagement and community building across collections, which previously often worked independently from one another or with a smaller group of institutions and/or collaborators. The attempt to create ever-growing biodiversity data aggregators to improve global research access to digital biodiversity data has made huge progress over the past decade and has resulted in increased availability of biodiversity data from fewer, larger data stores. It has also motivated unselfish collaboration between major aggregators in search of strategies for merging these data silos into a consolidated global data product. We describe an ongoing collaboration between the Global Biodiversity Information Facility ([GBIF](#)), The Atlas of Living Australia ([ALA](#)), Integrated Digitized

Biocollections (iDigBio), and the Distributed System of Scientific Collections ([DiSSCo](#)) to establish a global framework for integrating technologies, processes, standards, Application Programming Interfaces (APIs), ingestion, data, and data services, with the goal of building a well-documented linked system that relies on the various areas of expertise of the initial partners but with definitive pathways for incorporating new and existing entities as they desire or are developed.

We use the case of paleontological data as an exemplar of the potential impact of this collaboration. The iDigBio Paleontology Digitization Working Group, which was originally created by iDigBio as part of their community engagement program, has continued to be an active and engaged community of data providers and end-users, organizing numerous workshops and webinars. Currently, working group members, in collaboration with iDigBio staff and developers, are examining issues specific to paleontologic data aggregation that were identified by data providers; they are also working on a series of best-practices guidelines for sharing paleontologic data that will ideally help to reduce the number of mistakes made by downstream data aggregation manipulations. The focus of the working group is, and has been, largely community driven and supported by iDigBio through the provision of virtual meeting space for participants and by hosting the group's wiki-page of resources. Additionally, iDigBio has been proactive in working with other digitization initiatives in the paleontologic community (e.g., [Paleobiology Database](#)) on projects such as [ePANDDA](#) (enhancing Paleontological and Neontological Data Discovery API), which seeks to link existing digital resources through API development.

Keywords

global alignment, global aggregation, iDigBio, global data product

Presenting author

Gil Nelson

Presented at

Biodiversity_Next 2019

Funding program

iDigBio is funded by grants from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program [DBI-1115210 (2011-2018) and DBI-1547229 (2016-2021)]. Any opinions, findings, and conclusions or recommendations expressed in

this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Grant title

iDigBio Phase 2